DOCUMENT RESUME

ED 206 028 CS 503 566

AUTHOR Hample, Dale
TITLE Models of Arguments Using Multiple Bits of Evidence.
PUB DATE May 81
NOTE 30p.: Paper presented at the Annual Meeting of the International Communication Association (31st, Minneapolis, MN, May 21-25, 1981).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Cognitive Processes; College Students; *Communication Research; Comparative Analysis; *Credibility; Decision Making; Higher Education; *Models; *Persuasive Discourse; Validity

ABSTRACT

A study tested several different cognitive models of the way people use supporting arguments to arrive at belief in a claim. The models tested were (1) the model based on the assumption that human information processing is essentially logical; (2) R. S. Wyer's model; (3) the message-plus-context model; and (4) the weighted average model. One hundred fifty-six college students were administered a pretest, read one or two messages, and then were given a posttest. Each message contained three arguments supporting the same claim. The original outline for each message consisted of a series of three syllogisms, and the messages were written in such a way as to flesh out the premises. The results revealed theoretical and empirical support for the message-plus-context model, but comparably accurate predictions were also achieved from the Wyer and the weighted average models. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED206028

MODELS OF ARGUMENTS USING MULTIPLE BITS OF EVIDENCE

Dale Hample

Western Illinois University

Abstract

This study tests five different cognitive models of the way people use several supporting arguments to arrive at belief in a claim. Results point to the plausibility of all five models. Superior theoretical and empirical support weigh in favor of the message plus context models, but comparably accurate predictions follow from the Wyer and data-weighted average models. The models generally account for a quarter to a third of the variance in claim beliefs.

Presented to the annual meeting of the International Communication Association (Division I), Minneapolis, 1981.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Dale Hample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS503566

MODELS OF ARGUMENTS USING MULTIPLE BITS OF EVIDENCE

One way to understand argument is to view it as a cognitive event. Arguments are therefore comprised of beliefs, not propositions, and point toward decisions or attitudes, not message conclusions. Arguments are to be found in people's understandings, rather than in their conversations or essays. In short, this view insists that messages stimulate but do not contain arguments.

This paper continues a series of investigations intended to describe the process whereby people move cognitively from evidence to conclusion (Hample, 1977a; 1978; 1979a; in press). In particular, this study is designed to improve on the previous work in three ways: by including more than one bit of evidence in the justification of claims; by providing a more complete test of the logical assumptions underlying the basic model; and by comparing the model with some other competitive formulations. The details of and rationales for these improvements are given in the next section.

Theory

Belief processing can be modeled in quite a number of ways, and each procedure has its own theoretical presuppositions. The competing models to be tested are explained here.

The Basic Model

The model which has been used in the earlier studies is based on the assumption that human information processing is essentially logical. It is, of course, a well-established fact that people make many logical errors (see Hample, 1979b; 1980), but this does not mean that a logical model is useless: only that it will be imperfect, and that a search for exceptions might prove profitable. If our choice is between seeing people as logical reasoners who often make mistakes versus seeing people as mistakers who often are logical, the first alternative seems to have more

potential for successful modeling. At any rate, it is the choice made here.

The first step in constructing the model is to find a logical formulation which can encompass any kind of reasoning. A satisfactory solution is the hypothetical syllogism. It will accomodate any logical inference; it has, in fact, been argued to be the real basis of the Toulmin model (Hample, 1977b; see Toulmin, 1958). This syllogism insists that the movement from data (D) to claim (C) is mediated by the idea that if the data are true, the claim is true. Hence: 'if D is true, C is true; D is true; so C is true.' This basic relationship should also hold if D and C are allowed to be <u>probably</u> true or false as well. To see this, consider that C may be viewed as a joint event, which occurs only when D and 'if D, then C' (i.e., C|D) also occur (see Hays, 1973, for a guide to the following development). Representing C|D for convenience as W (for warrant), we have

(1) $C = D \cap W$.

Assuming that D and W are logically independent, then we may also say that

(2) $p(C) = p(D \cap W) = p(D)\ p(W)$, or

(3) $p(C) = p(D)\ p(C|D)$.

Assuming that (3) applies to subjective probabilities as well as formal mathematical ones, we may interpret it this way: a person will believe in C to the degree that he/she believes that D and C|D are both likely.

Equation (3) is the basic model for the rest of the paper, though it will be elaborated and applied in different ways. Conclusion drawing is seen as a logical joint event which is causally dependent on the subjective probabilities of both data and warrant.

But equation (3) is not expected to be sufficient as it stands. If a person were to approach a new argument tabula rasa, the basic formulation would be fine. But in real-world arguments, people test new information against their context for C (see Hample, in press). A given individual might well have high probabilities for D and C|D, but still be disinclined to think C true because of her/his previous

beliefs. This context contains the results of all earlier arguments regarding C. McGuire (1960) noticed this problem, but was unable to solve it usefully because he could not specify the "other" arguments (i.e., the context) which bear on C. Wyer (1974; Wyer & Goldberg, 1970), however, found a way around the apparent need to identify the "other" beliefs. He suggests use of another clause: $p(\bar{D})\ p(C|\bar{D})$. $\bar{D}$ (not D) refers to the probability that D is false, and $C|\bar{D}$ envisions the likelihood that C may be true, even given that D is not. So we might way

(4) $p(C) = p(\bar{D})\ p(C|\bar{D})$

in order to emphasize that "other" arguments--those not dependent on D being true--may still influence C. Notice that (4) is a version of (3) which uses $\bar{D}$ instead of D as the first premise.

Wyer's Model

Actually, Wyer's model did not itself develop in this way. He obtained it from part of the basis for Bayes' Theorem (see Winkler, 1972, ch. 2, for the following). Suppose we are trying to explain the dependency of B on A. We may distinguish between two kinds of B: one which intersects with A ($B \cap A$) and one kind which does not ($B \cap \bar{A}$). The two types of B are, of course, mutually exclusive. This allows us to assert that p(B) will be the sum of the probabilities of the two classes of B. That is,

(5) $p(B) = p(A \cap B) + p(\bar{A} \cap B)$.

Both right members of (5) are joint events. In situations like the one described here, the joint probability of two events E_1 and E_2 is

(6) $p(E_1 \cap E_2) = p(E_1)\ p(E_2|E_1)$.

Applying (6) to (5) gives us

(7) $p(B) = p(A)\ p(B|A) + p(\bar{A})\ p(B|\bar{A})$,

which is Wyer's model for inferring B on the basis of A. It is simple enough to make (7) into a model of argument; one need only notice that B is a conclusion based on A. Therefore,

(8) $p(C) = p(D)\ p(C|D) + p(\bar{D})\ p(C|\bar{D})$.

Equation (8) combines the concerns of (3) and (4) into a single prediction. (This development assumes that neither [3] nor [4] is a fully accurate prediction of p[C]; otherwise, one could conclude that p[C] = p[C] + p[C] from [8].)

The Message Plus Context Model

The discussion of equations (3) and (4) interpreted them quite differently. The right hand side of (3)--p(D) p(C|D)--describes the "new" syllogism which might affect p(C), while equation (4) deals with the old syllogisms, or context for p(C). In a persuasive situation, the message will try to stimulate a high value for p(D); this means that stimulation of a change in p(C) is attempted via p(D) p(C|D). Although (8) allows for both the message-stimulated syllogism and the cognitive context to affect p(C), (8) insists that they have equal weight in doing so. For either half of the right hand predictor to have extra weight would violate the derivation of the model. Wyer's development does not support

(9) $p(C) = \beta_1\ p(D)\ p(C|D) + \beta_1\ p(\bar{D})\ p(C|\bar{D})$,

which allows for differential weighting of message and context syllogisms.

Fortunately, there is another way to arrive at (9). If we temporarily ignore the relation between the two predictors, we have a familiar situation--one variable being predicted by two others:

(10) $p(C) = \beta_1(\text{Message}) + \beta_2(\text{Context})$.

The two predictor syllogisms are defined by equations (3) and (4), respectively. When these predictions are substituted into (10), we are left with (9).

This line of reasoning encourages us to give attention to the relationship between p(D) p(C|D) and $p(\bar{D})\ p(C|\bar{D})$. If (8) is essentially accurate, then (9) should yield roughly equivalent weights for the two predictors, and equations (8) and (9) should be equally effective in predicting p(C). In fact, (9) can reduce to (8) in the case where $\beta_1 = \beta_2 = 1$. This result would sustain the assumptions

of the Wyer development, and make the message plus context conceptualization unnecessary.

Yet another possibility deserves attention. In (5) we assumed that the two right hand probabilities were mutually exclusive of one another. In the imperfect world of subjective probabilities, this might not be the case. If we were to suppose that message and context syllogisms were not necessarily mutually exclusive, we would have to reconsider (9). Both (9) and (10) regard p(C) as the union of message and context syllogism--i.e.,

(11) $p(C) = p(\text{Message} \cup \text{Context})$,

with suitable weights added to modify this logical model. If message and context are not mutually exclusive, (11) expands not to (10), but to (12):

(12) $p(C) = p(\text{Message}) + p(\text{Context}) - p(\text{Message} \cap \text{Context})$,

from which we obtain

(13) $p(C) = p(\text{Message}) + p(\text{Context}) - p(\text{Message})\ p(\text{Context})$.

Substitutions and the insertion of beta weights results in

(14) $p(C) = \beta_1 p(D)\ p(C|D) + \beta_2 p(\bar{D})\ p(C|\bar{D}) + \beta_3 p(D)\ p(C|D)\ p(\bar{D})\ p(C|\bar{D})$.

Significance testing of β_3 can determine whether the third predictor has any use; if not, the mutual exclusivity assumption will have been supported.

The Weighted Average Models

Anderson (1971; 1974; 1979) has shown that people consistently seem to use fairly simple "cognitive algebras" in combining information. He has found particularly good support for a weighting formulation:

(15) $R = \Sigma(w_i\ s_i)/\Sigma w_i$

where R is a general response of some kind, s_i are the scale values of the i information relevant to R, and the w_i are the weights assigned to each s. Anderson's whole model is much more complex than indicated in (15), and he insists on a particular means of evaluating the model--a means which is not used here, since it would not be applicable

to all the tested models. All that we take from (15) is the idea of using a weighted average to predict R.

Part of the application of (15) to our problem is obvious: R is p(C). But shall we predict p(C) with data weighted by warrants, or with warrants weighted by data? Although (3) emphasizes that the functions of data and warrant are clearly distinct, we should not assume the data and warrant beliefs take clearly distinguishable forms (see Hample, 1977b). In Toulmin's (1958) classic example, he proves that 'Harry is a British citizen.' To support this claim, Toulmin says that 'people born in Bermuda are British citizens' and that 'Harry was born in Bermuda.' Which of these is more evidential? In fact, both seem to make data-like assertions, and either could warrant the link between the other and the claim. So we need a version of (15) for both possibilities:

(16) $p(C) = \Sigma(p[D]\ p[C|D])/\Sigma p(D)$

(17) $p(C) = \Sigma(p[D]\ p[C|D])/\Sigma p(C|D)$.

Equations (16) and (17) suffer, however, from relying only on the message syllogism to predict p(C). To take into account the cognitive context requires only minor alteration:

(18) $p(C) = \Sigma(p[X]\ p[Z])/\Sigma p(X)$

(19) $p(C) = \Sigma(p[X]\ p[Z])/\Sigma p(Z)$,

where X represents both D and $\bar{D}$, and Z stands in for both $C|D$ and $C|\bar{D}$--whichever is appropriate for the X.

Multiple Evidence (The Hypotheses)

All the preceding equations have shared one simplifying assumption (one which characterizes nearly all our theorizing about argument). They all cover the case where a message stimulates processing of a single D. But as Becker (1978) has so usefully shown, our beliefs are often influenced by multiple bits of evidence. These may appear at more or less the same time, or some event may suddenly stimulate

a person to process a number of items which had been previously noted and remembered, but never really integrated into the belief system. To be realistic, we need argument models which explain how p(C) is affected by multiple Ds. By making the assumption that arguments using different Ds are independent and mutually exclusive of one another, straightforward extensions of the various models are possible.

$$(20)\quad p(C) = \Sigma[p(D)\ p(C|D) + p(\bar{D})\ p(C|\bar{D})]$$

$$(21)\quad p(C) = \beta_1\ p(D)\ p(C|D) + \beta_2\ p(\bar{D})\ p(C|\bar{D}) + \beta_3\ p(D)\ p(C|D) + \ldots$$

$$(22)\quad p(C) = \beta_1\ p(D)\ p(C|D) + \beta_2\ p(\bar{D})\ p(C|\bar{D}) + \beta_3\ p(D)\ p(C|D)\ p(\bar{D})\ p(C|\bar{D}) + \beta_4\ p(D)\ p(C|D) + \ldots$$

$$(23)\quad p(C) = \Sigma(p[X]\ p[Z])/\ \Sigma p(X)$$

$$(24)\quad p(C) = \Sigma(p[X]\ p[Z])/\ \Sigma p(Z)$$

Equations (20) through (24) constitute the hypotheses of this study. Equation (20) is the basic Wyer formulation, while (21) and (22) are the message plus context models, with and without the assumption that message and context are mutually exclusive. The final two equations are the versions of the weighted average model. All five equations make the simplifying assumption that each argument has its own context; it is possible to describe the common context, but little harm should be done by this simplification.

The equations will be tested with correlational methods. Pearson's r will be used for (20), (23) and (24), and multiple regression for (21) and (22).

It only remains to make explicit how this investigation differs from previous ones. There are three main distinctions: (1) earlier studies (except Hample, 1977a) have used only a single stimulated argument, and the exception does not offer a thorough analysis of the multiple evidence problem; (2) the possibility that the message and context syllogisms might not be mutually exclusive has never been tested, as it will be here in the comparison of (21) and (22); and (3) the weighted average formulations have not been tried before.

Method

Subjects were undergraduates earning extra credit in a university-required public speaking course. Administration of the study occurred out of class in groups of about 10 to 30. The study required participation in two sessions a week apart. Data were provided by 173 subjects, of whom 25 did not return for their second session. Two subjects were excluded for failure to follow directions, and one student exercised his/her option not to respond at all. Subjects' names were used to match responses from the two sessions together. Students were debriefed after the second session, at the end of which they completed an open-ended questionnaire eliciting their reactions to the study. No one came close to figuring out the hypotheses, but a large number--perhaps a third--spontaneously remarked that the task was very boring.

At the first session, subjects filled out a pretest, read one or two messages, and filled out a posttest. At the second session, students completed the delayed posttest, which was identical to the pretest. The first session usually required 30 to 40 minutes, and the second was about half as long.

All of the test items had the same general form: a statement would be given, and subjects would rate it on a seven-step LIKELY-UNLIKELY scale. Items were of several specific types. The most common form provided a simple declarative statement--e.g., "The job market for new college graduates is very poor"--which was followed by the subjective probability (i.e., likelihood) scale. This type of item measured $p(C)$, $p(D)$ and $p(W)$. Inserting a negative into such statements yielded the measures of $p(\bar{D})$ and $p(\bar{W})$. Actually, warrants were stated in two ways: once simply (e.g., "Teaching jobs are important in the labor market for new college grads") and once conditionally (e.g., "Assume it is TRUE that few teaching jobs are available now. With that in mind, rate: The job market for new college graduates is very poor"). The first type of item gives $p(W)$ or $p(\bar{W})$. The second type yields measures of $p(C|D)$ or $p(C|\bar{D})$, depending on whether subjects assumed

something to be true or false. By assuming three things to be false at once, subjects also estimated $p(C|\bar{D}\bar{D}\bar{D})$, but these measures proved too unreliable to use in analysis. Items were distributed randomly throughout the booklets. Students responded on sense-mark computer forms. These various procedures generated all the measures required by equations (20) through (24), and allow each to be tested with p(W) as well as p(C|D), or with $p(\bar{W})$ substituted for $p(C|\bar{D})$. The pretest and delayed posttest were prefaced by nine items of similar forms, but which were unrelated to the message topics or each other. These items were included to allow an opportunity for students to get used to the rating procedures, and to provide test-retest reliability estimates. The pretest and delayed posttest had a total of 49 items, and the immediate posttest had 40.

The experimental items were concerned with two messages. One--on the job market for college graduates--has been used before (Hample, 1977a). The other--on a tax cut--was written for this study. Both messages contained three arguments supporting the same claim. The original outline for each message consisted of a series of three syllogisms, and the messages were written in such a was as to flesh out the premises. No effort was made to use logically correct phrasing. The job market message was represented as an extract from a radio interview show, and the tax cut speech was described as part of a speech given by a Congressional candidate. Each was about 750 words long. Every student read either the job speech, the tax speech, or both. These conditions were assigned randomly. Therefore, about one third of the sample is a control group for each message.

Results

Reliability

The nine warm-up items appeared on the pretest and the delayed posttest, one week later. Five of these items were of the simple p(A) type, three were of the conditional p(B|A) type, and one was conditional, with three things assumed

false $p(B|\overline{AAA})$. Test-retest correlations were calculated, converted to z scores, and averaged. Results for all possible subjects appear in the first column of Table 1. Reliabilities are low, especially for the conditional items. Cohen (1978),

INSERT TABLE 1

using 100-point scales in a subjective probability study of trait ratings, reports average correlations of about .65 for the p(A) type of item, about .60 for the $p(B|A)$ type, and about .40 for the $p(B|\overline{AAA})$ type. Scale and item topic differences prevent exact comparisons between Cohen's results and the present ones, but reliabilities seem generally comparable for the p(A) items.

The low reliabilities, combined with the post-experimental complaints about boredom, encourage the belief that some subjects did not take their task to heart. An effort was therefore made to reduce the data by excluding subjects who make unreliable responses. Scores from all nine warm-up items were added together, and the delayed posttest sum was subtracted from the pretest sum. The mean difference was -1.685, with a standard deviation of 4.431. For the first reduction, all subjects were excluded whose difference score was not between 4 and -7 (approximately $\pm$ 1 s). This reduced sample size to about 95 (the procedure also excludes every subject with any missing data on these items). Reliabilities improved noticeably, as the second column of Table 1 shows. An even more restrictive reduction was tried, and subjects' difference scores were required to be between 3 and -6. This second reduction excludes about 15 more subjects, as Table 1 shows, and improves the $p(B|A)$ reliabilities. But there is an obvious trade-off between reliability on one hand and statistical power (i.e., sample size) on the other. The first reduction was judged to yield the best balance, and so further analyses were conducted with these subjects. The intention of the reduction procedure is simply to eliminate bored, unreliable subjects. Because of the possibility that the exclusion technique

is distorting results in some other way, however, all results will be reported for both "all" and "reduced" data sets throughout this paper.

Validity

Equation (9) has been applied to similar data in earlier studies of one argument messages (Hample, 1978, 1979a, in press). The usual result in those studies has been that (9) explains between half and a quarter of the variance in p(C). Equation (9) was applied to each of the three arguments stimulated by each message at each time. Table 2 reports the results, which are the mean multiple

---

INSERT TABLE 2

---

correlations (averaged using z transformations) for three arguments on two topics. Each $\bar{R}$ in Table 2, therefore, represents six multiple correlations. Results were categorized according to testing time (pretest, immediate or delayed posttest), whether W or a conditional statement was used to measure the warranting function, and whether subjects read the pertinent message ("experimental") or not ("control").

Pretest results using (9) have usually been disappointing, but those for the present reduced data are solid. In contrast, the immediate posttest results for the reduced sample are lower than usual. In general, however, the results conform to those expected on the basis of previous work; this comparability is evidence for the data's construct validity. Table 2 also indicates that the C|D formulation of the warrant produces more accurate estimates of p(C) than does the simpler W phrasing, and that experimental results are more accurate than control data. The final important result in Table 2 concerns the Socratic Effect (McGuire, 1960; Wyer, 1974), which is the robust finding that logical consistency increases after people have had time to integrate related information given to them either by questionnaire or message. This effect occurs only between the two posttests for the reduced sample, and between the pretest and immediate posttest for all

subjects. But taken as a whole, the pattern of results in Table 2 is evidence of satisfactory validity for the data, especially when the conditional warrrant forms are used.

Persuasion

Although the hypotheses do not depend on the assumption that subjects are changing their beliefs, if beliefs are in flux during the study, the hypotheses get a more varied test. For that reason, the study was modeled after the usual attitude change paradigm. Comparisons were calculated in order to determine whether or not belief change occurred as a result of exposure to the message stimuli. Within the reduced sample, correlated t-tests show significant change in p(C) from the pretest to either the immediate or delayed posttest on both topics for the experimental groups ($p<.001$). Only for the pretest to immediate posttest on the tax topic does this effect occur for the control group ($p<.05$). Direct comparisons of the experimental versus control groups show no significant differences on the pretest, but statistically significant differences on both posttests ($p<.05$). For the reduced data, therefore, experimental subjects were persuaded and the control subjects generally were not. For the whole data set, the experimental subjects showed significant change from the pretest to both posttests ($p<.05$), except for the tax topic's pretest to immediate posttest comparison. Control subjects did not alter their attitudes at significant levels ($p>.05$). Direct comparisons of the experimental and control subjects, however, failed to produce significant differences ($p>.05$). In sum, persuasion seems to have occurred clearly in the reduced data set, but only weakly in the total sample.

Hypotheses

Results bearing on the hypotheses will first be discussed one hypothesis at a time, and then overall comparisons will be made.

The Wyer equation (20). Equation (20) specifies that the unweighted predictors, $p(D)\ p(C|D)$ and $p(\bar{D})\ p(C|\bar{D})$, be calculated for each of the three arguments,

then added together to predict p(C). Alternatively, W may be substituted for C|D, and $\bar{W}$ for C|$\bar{D}$. Results are reported in Table 3. As is immediately apparent from

INSERT TABLE 3

inspection of the table, this equation does a good job of fitting the data. Correlations are better than those for the validity results in Table 2, and are particularly strong for the C|D conditions for the reduced data. Equation (20) is therefore supported.

Message plus exclusive context (21). Table 4 indicates the results of predictions of p(C) based upon each of the six message or context syllogisms being treated as a separate variable. Predictions are good again, particularly when C|D

INSERT TABLE 4

operationalizes the warrants. When W is used, correlation coefficients tend to be about .10 lower than when C|D appears. Even the W-based predictions, however, explain about a quarter of the total variance in p(C).

Message plus nonexclusive context (22). The difference between equations (21) and (22), the reader will recall, is that (22) drops the assumption that the message and context syllogisms are mutually exclusive. Each argument therefore yields an additional term (the product of the message and context syllogisms) to be tested. The essential question is whether or not the additional term is required by the data. Table 5 reports results which address that issue. To obtain the

INSERT TABLE 5

multiple correlations reported in Table 5, the message and context phrases of (22) were entered into the regression equation; after the variance explained by those terms had been accounted for, the product terms were entered. Significance tests on the product terms therefore precisely confront the issue of whether the product

terms are important once the context and message syllogisms are considered. In addition to reporting multiple correlations, Table 5 indicates the number of times the product term is significant ($p<.05$) for each prediction. Since each prediction could have had three significant products (because there are three arguments for each predicted p[C]), the maximum possible for the whole table is 144. The product actually is significant 13 times, or about 9% of the time. This is more than the 5% expected by chance ($z=2.21$, $p<.05$). The reader may note that the multiple correlation coefficients are often not significant in Table 5. When only the 24 significant predictions are considered, the products are in turn significant in 7 of 72 cases, or about 10% of the time ($z=5.43$, $p<.001$). Multiple regressions were also run on each argument separately (as was done to obtain the validity data), using message, context and their interactions as variables. For these regressions, the products are significant 17% of the time, and 23% of the times R is itself statistically significant. The proportions are more than chance would predict, but are still not strikingly high ($z=6.79$, $p<.001$, and $z=8.07$, $p<.001$, respectively). The evidence in favor of including the product term is therefore weak but statistically significant. It should also be noted that the overall predictions in Table 5 are generally better than those in Table 4.

Data-weighted averages (23). Equation (23) is the first of the weighted average models, and specifies that the sum of the six message and context syllogisms be divided by the sum of the six values of p(D) and $p(\bar{D})$ in order to predict p(C). Table 6 reports results. These correlations are fairly high, and are most impressive

---

INSERT TABLE 6

---

for the reduced C|D data. This equation produces results generally comparable to those for equations (20), (21) and (22). Like them, it is supported.

Warrant-weighted averages (24). The final hypothesized model is equation (24),

which differs from (23) only in that the sum of W and $\bar{W}$ (or C|D and C|$\bar{D}$) is used as the denominator. When W is used in the denominator, it is also used to calculate

INSERT TABLE 7

the numerator, and similarly for C|D. Results are in Table 7. Predictions are generally inferior to those from Table 6, especially for the C|D data in both the reduced and full sets. However, the correlation coefficients are still high enough to provide some support for equation (24), though of course the previous models are clearly better.

Overall comparison. Table 8 reports average correlations (calculated using the r-to-z transformation) for the various hypotheses. Results from the validity check are also provided for comparison. The data show that the Wyer model (equation

INSERT TABLE 8

20), both message plus context equations (21 and 22) and the data-weighted average formula (23) do adequate and comparable jobs. The best overall model is the non-exclusive message plus context equation (22), but it should be noted that much of the advantage can be traced to the W-based predictions. Table 8 shows that C|D is generally the better operationalization of warranting, and therefore those results ought to be weighed more heavily. For the best formulations (equations 20, 21, 22 and 23, using C|D), the experimental predictions are usually much better than those for the control groups. The reduced data set also produced better predictions, when compared to the full set; however, the reduction had very little effect on the study's general pattern of results. Finally, Table 8 shows that the better models explain between a quarter and a third of the variance in p(C).

Discussion

Though one might well wish for data of higher quality, the reliability and

validity of these students' responses are sufficient to meet this study's objectives.

The study's main goal has been to model arguments which use more than one bit of evidence. Examination of Table 8 shows that this can be done in several acceptable ways. First, the researcher might make separate predictions of p(C) for each supporting argument, and average the predictions (this is the technique involved in the validity estimates). This is not the best way, however, and there are good theoretical reasons for believing that it will never be; the procedure does not take account of the relationships among the supporting arguments, and makes the implicit assumption that each argument should have equal weight in the prediction. This assumption is arbitrary, and there is no reason to suppose it will ordinarily be accurate.

Next best is the Wyer model. With single arguments, the Wyer formula has been worse than the multiple regressions which were averaged for the validity data (Hample, 1978; 1979a; in press). But here the Wyer model is better. This shows the weakness of the validity procedure, which is weighted within but not between arguments. These results also suggest the possibility that the Wyer model may be increasingly accurate with larger numbers of supportive arguments.

Slightly better jobs are done by the exclusive message plus context and data-weighted average models. Equation (21) is a straightforward extension of the basic model, and results continue to display support for the view that arguments consist of data and warrant. That overall predictions here tend to be lower than in previous studies may be attributed to the admittedly marginal quality of the data. Equation (23) also produced good predictions. The advantage to this formula is that it does not require separate beta weights for each syllogism, and so conserves degrees of freedom. When results from equations (23) and (24) are compared, the superiority of data weighting is evident.

Though equation (23) is not completely consistent with the basic theory, results suggest that the data-weighted average may be a convenient computational

substitute for (21) or (22). However, a strong theoretical basis is an essential attribute for a model to possess, and (23) has the character of a theory-independent ad hoc modification. Both (21) and (22) have much better theoretical grounding. The best overall job of prediction is done by (22). This equation produces uniformly good predictions, even under the adverse conditions of using W and/or the full data set. Taking both theoretical and empirical support into consideration, the three best models (in order) for this study are the nonexclusive message plus context model, the exclusive message plus context model, and the data-weighted average model.

Another of the study's goals is to choose between some alternative procedures for modeling argument. On one count--whether W or C|D should be used to measure warranting--results are clearcut. Virtually without exception (see Table 8), the C|D predictions were better. The reason for this is probably that C|D has a more functional form. Data, warrant and claim are not viewed here as elements of a message; they are seen as cognitive functions which must occur whenever arguing occurs. The data and claim functions are fairly distinct, but the warranting function is not. People may well vary in the way they move from data to claim. In writing Ws for the questionnaires, it is necessary to anticipate the exact thought which converts belief D into belief C, and this is obviously difficult to do. The C|D items, on the other hand, always take the general form 'if D, then C.' Each person supplies his/her own reason (consciously or not) for believing C|D. The C|D form therefore captures the function of warranting without the need to specify the warrant. Certainly C|D makes better theoretical sense than W, and empirical results support the theory. This theoretical advantage seems to outweigh the fact that p(W) items will generally have higher reliabilities than conditional items.

The study's last goal is to test the assumption that the message and context syllogisms are mutually exclusive. Here the data are not completely unequivocal, but point toward nonexclusivity. The nonexclusive model (22) produces the better predictions. However, given the nature of multiple regression procedures, this is nearly inevitable. Even adding a random independent variable to a regression equation could never lower

R, and chance being what it is, R would sometimes be larger with the random variate. The added variables here were constructed from others which have clear relationships with p(C), and so small increases in R provide no evidence for the importance of the new variable. However, the increases in R are not always small, as Table 8 shows. The total mean R increases by about .075 in the reduced data, and .050 in the whole sample. For the control conditions, the increase is about .090. In the context of predictions whose Rs rarely exceed .70, these are important increments. More direct evidence on the importance of the message times context term appears in Table 5. The extra terms acheived statistical signficance about 10% of the time--significantly more than chance alone dictates. Here, then, is support for the inclusion of the new term. However, results do indicate that adequate predictions are possible with or without the assumption of exclusivity. The extra variables require additional degrees of freedom, and as comparison of Tables 4 and 5 shows, therefore have a tendency to lower the significance levels of R. With moderate sample size and a large number of arguments to test, pragmatic considerations of statistical power might well justify dropping the extra term. This study's overall judgment is that the additional terms implied by rejection of the mutual exclusivity assumption have a reasonable degree of theoretical and empirical support.

Conclusions

This study supports several conclusions; 1) The basic model of argument (equation 3) continues to display validity. 2) The Wyer model yields acceptable predictions of p(C) when multiple arguments are considered. 3) The data-weighted average model predicts p(C) well, but the warrant-weighted average model does not. 4) The message plus context models both perform well with multiple arguments; the nonexclusive version is best of all. 5) The warranting function is measured better with the conditional C|D phrasing than with the less function W form. 6) The context and message components of arguments are probably not mutually exclusive; however, little harm will normally be done by the simplifying assumption of exclusivity.

References

ANDERSON, N.H. Integration theory and attitude change. Psychologial Review, 1971, 78, 171-206.

ANDERSON, N.H. Cognitive algebra: Integration theory applied to social attribution. In L. Berkowitz (Ed.), Advances in experimental social psychology, vol. 7. New York: Academic Press, 1974, 1-101.

ANDERSON, N.H. Algebraic rules in psychological measurement. American Scientist, 1979, 67, 555-563.

BECKER, S. L. Visual stimuli and the construction of meaning. In B.S. Randhawa & W.E. Coffman (Eds.), Visual learning, thinking, and communication. New York: Academic Press, 1978, 39-60.

COHEN, J.L. Test-retest reliabilities of subjective probability judgments. Journal of Personality and Social Psychology, 1978, 36, 643-646.

HAMPLE, D. Testing a model of value argument and evidence. Communication Monographs, 1977a, 44, 106-120.

HAMPLE, D. The Toulmin model and the syllogism. Journal of the American Forensic Association, 1977b, 14, 1-9.

HAMPLE, D. Predicting immediate belief change and adherence to argument claims. Communication Monographs, 1978, 45, 219-228.

HAMPLE, D. Predicting belief and belief change using a cognitive theory of argument and evidence. Communication Monographs, 1979a, 46, 142-146.

HAMPLE, D. A review of empirical literature on logical processes. Paper presented to Speech Communication Association, San Antonio, 1979b.

HAMPLE, D. Symbolization and fallacy. Paper presented to Speech Communication Association, New York City, 1980.

HAMPLE, D. The cognitive context of argument. Western Speech Communication Journal, in press.

HAYS, W.L. Statistics for the social sciences. New York: Holt, Rinehart & Winston, 1973.

McGUIRE, W.J. A syllogistic analysis of cognitive relationships. In C.I. Hovland & M.J. Rosenberg (Eds.), Attitude organization and change. New Haven: Yale University Press, 1960, 65-111.

TOULMIN, S.E. The uses of argument. Cambridge: Cambridge University Press, 1958.

WINKLER, R.L. Introduction to Bayesian inference and decision. New York: Holt, Rinehart & Winston, 1972.

WYER, R.S., Jr. Cognitive organization and change: An information processing approach. Potomac, Md.: Lawrence Erlbaum, 1974.

WYER, R.S., Jr. & GOLDBERG, L. A probabilistic analysis of the relationships among beliefs and attitudes. Psychological Review, 1970, 77, 100-120.

TABLE 1

Test-Retest Reliabilities (Average Correlations)

Item Type	all Ss[1]	first reduction[2]	second reduction[3]
$p(A)$	.645	.691	.693
$p(B\|A)$	.357	.491	.549
$p(B\|\bar{A}\bar{A}\bar{A})$	.170	.247	.218

[1] N is about 145. All original correlations were significant at $p<.001$, except $p(B|\bar{A}\bar{A}\bar{A})$, where $p<.05$.

[2] N is about 95. All original correlations were significant at $p<.001$, except $p(B|\bar{A}\bar{A}\bar{A})$, where $p<.01$.

[3] N=79. All original correlations were significant at $p<.001$, except $p(B|\bar{A}\bar{A}\bar{A})$, where $p<.05$, and one of the $p(B|A)$, where $p<.01$.

TABLE 2

Validities (Average Multiple Correlations Predicting p(C) with One Argument)

Test Time:		1	2	3
Condition	Warrant			
		Reduced Data		
Exper.	C|D	.690 (56)	.495 (56)	.611 (56)
	W	.583 (56)	.362 (56)	.510 (56)
Control	C|D	.600 (30)	.394 (30)	.565 (30)
	W	.482 (30)	.280 (30)	.534 (30)
		All Data		
Exper.	C|D	.408 (107)	.586 (106)	.525 (91)
	W	.316 (107)	.565 (106)	.468 (91)
Control	C|D	.387 (53)	.582 (53)	.495 (45)
	W	.366 (53)	.412 (53)	.382 (45)

Note. Average sample size is in parentheses.

TABLE 3

Correlations Between p(C) and the Wyer Predictor

Test Time:			1	2	3
Condition	Topic	Warrant			
			Reduced Data		
Exper.	Tax	C|D	.526***(54)	.780***(53)	.795***(57)
		W	.272*(54)	.568***(53)	.527***(58)
	Job	C|D	.440***(60)	.697***(60)	.652***(61)
		W	.066 (62)	.551***(61)	.622***(60)
Control	Tax	C|D	.527***(36)	.795***(35)	.642***(34)
		W	.262†(36)	.568***(35)	.653***(32)
	Job	C|D	.541***(29)	.744***(27)	.472**(29)
		W	.268†(28)	.530**(28)	.063 (29)
			All Data		
Exper.	Tax	C|D	.415***(111)	.661***(104)	.632***(91)
		W	.317***(114)	.552***(104)	.514***(92)
	Job	C|D	.661***(104)	.645***(114)	.539***(94)
		W	.552***(104)	.541***(114)	.515***(93)
Control	Tax	C|D	.399***(60)	.750***(58)	.594***(48)
		W	.403***(59)	.569***(58)	.318*(47)
	Job	C|D	.425***(49)	.675***(49)	.517***(47)
		W	.259*(51)	.405***(50)	.174 (47)

Note. Sample sizes in parentheses.

†p<.10

*p<.05

**p<.01

***p<.001

TABLE 4

Multiple Correlations Between p(C) and the Message Plus Exclusive Context Model

Test Time:			1	2	3
Condition	Topic	Warrant			
			Reduced Data		
Exper.	Tax	C|D	.525**(54)	.667***(53)	.633***(57)
		W	.462 (54)	.515*(53)	.578**(58)
	Job	C|D	.753***(60)	.677***(60)	.719***(61)
		W	.483*(62)	.524**(61)	.712***(60)
Control	Tax	C|D	.485 (36)	.504 (35)	.553 (34)
		W	.356 (36)	.471 (36)	.465 (32)
	Job	C|D	.436 (29)	.651 (27)	.608 (29)
		W	.643 (28)	.378 (28)	.312 (29)
			All Data		
Exper.	Tax	C|D	.530***(76)	.610***(76)	.610***(76)
		W	.380†(76)	.525***(76)	.563***(76)
	Job	C|D	.623***(86)	.705***(86)	.686***(86)
		W	.402*(86)	.634***(86)	.600***(86)
Control	Tax	C|D	.563*(43)	.488(43)	.593*(43)
		W	.414 (43)	.429 (43)	.468 (43)
	Job	C|D	.565†(37)	.646**(37)	.526 (37)
		W	.719***(37)	.307 (37)	.401 (37)

Note. Sample sizes in parentheses.

†p<.10

*p<.05

**p<.01

***p<.001

TABLE 5

Multiple Correlations Between p(C) and the Message Plus Nonexclusive Context Model, with Frequency of Significant Products Noted

Test Time:			1	2	3
Condition	Topic	Warrant			
			Reduced Data		
Exper.	Tax	C|D	.563*(0)	.706***(1)	.646***(0)
		W	.511 (0)	.547†(0)	.594**(0)
	Job	C|D	.786***(1)	.693***(0)	.719***(0)
		W	.503†(0)	.537*(0)	.730***(0)
Control	Tax	C|D	.552 (0)	.581 (0)	.586 (0)
		W	.465 (0)	.652†(1)	.638 (1)
	Job	C|D	.537 (0)	.791*(1)	.689 (0)
		W	.698 (0)	.497 (0)	.564 (1)
			All Data		
Exper.	Tax	C|D	.539**(0)	.639***(0)	.612***(0)
		W	.446†(0)	.562**(0)	.571***(0)
	Job	C|D	.646***(0)	.719***(0)	.720***(1)
		W	.416†(0)	.640***(0)	.629***(0)
Control	Tax	C|D	.609†(0)	.504 (0)	.640*(0)
		W	.580†(1)	.542 (0)	.598†(1)
	Job	C|D	.728**(2)	.773***(1)	.632†(1)
		W	.732**(0)	.347 (0)	.513 (0)

Note. Sample sizes may be found in Table 4.

†p<.10

*p<.05

**p<.01

***p<.001

TABLE 6

Correlations Between p(C) and the Data-Weighted Average Model

Test Time:			1	2	3
Condition	Topic	Warrant			
			Reduced Data		
Exper.	Tax	C|D	.605***(54)	.806***(53)	.817***(57)
		W	.306*(54)	.577***(53)	.509***(58)
	Job	C|D	.489***(60)	.752***(59)	.613***(61)
		W	.175†(60)	.535***(60)	.567***(60)
Control	Tax	C|D	.602***(36)	.824***(35)	.630***(34)
		W	.369*(36)	.530***(36)	.661***(32)
	Job	C|D	.501**(29)	.683***(27)	.567***(29)
		W	.202 (28)	.440**(28)	.064 (29)
			All Data		
Exper.	Tax	C|D	.590***(98)	.633***(103)	.653***(91)
		W	.304***(98)	.606***(103)	.511***(92)
	Job	C|D	.425***(111)	.675***(113)	.509***(94)
		W	.348***(112)	.531***(113)	.430***(93)
Control	Tax	C|D	.499***(60)	.805***(58)	.646***(47)
		W	.486***(59)	.558***(59)	.086 (46)
	Job	C|D	.358**(49)	.615***(49)	.586***(47)
		W	.337**(50)	.298*(50)	.185 (47)

Note. Sample sizes in parentheses.

†p<.10

*p<.05

**p<.01

***p<.001

TABLE 7

Correlations Between p(C) and the Warrant-Weighted Model

Test Time:			1	2	3
Condition	Topic	Warrant			
			Reduced Data		
Exper.	Tax	C|D	.153 (52)	.564***(53)	.504***(57)
		W	.142 (54)	.607***(53)	.552***(58)
	Job	C|D	.389***(60)	.637***(60)	.582***(61)
		W	-.122 (62)	.552***(61)	.681***(60)
Control	Tax	C|D	.131 (36)	.577***(35)	.518***(34)
		W	.375*(36)	.522***(36)	.643***(32)
	Job	C|D	.424*(29)	.605***(27)	.207 (29)
		W	.334*(28)	.611***(28)	.164 (29)
			All Data		
Exper.	Tax	C|D	.102 (96)	.439***(103)	.382***(91)
		W	.124 (98)	.617***(103)	.461***(92)
	Job	C|D	.360***(111)	.624***(113)	.583***(93)
		W	.033 (114)	.562***(114)	.664***(92)
Control	Tax	C|D	-.067 (60)	.543***(58)	.328*(48)
		W	.394***(59)	.519***(59)	.365**(46)
	Job	C|D	.482***(48)	.566***(49)	.292*(47)
		W	.282*(51)	.436***(50)	.270*(47)

Note. Sample sizes in parentheses.

*p<.05

**p<.01

***p<.001

TABLE 8

Average Zero-Order or Multiple Correlations

Warrant:		W		C|D		
Condition:		Exper.	Control	Exper.	Control	TOTAL
Equation	Table					
			Reduced Data			
[validity]	2	.490	.438	.605	.525	.517
20	3	.452	.411	.672	.636	.551
21	4	.552	.445	.668	.544	.557
22	5	.577	.592	.691	.633	.625
23	6	.451	.396	.699	.648	.563
24	7	.435	.457	.506	.426	.457
			All Data			
[validity]	2	.455	.387	.510	.492	.462
20	3	.502	.362	.599	.574	.515
21	4	.524	.469	.631	.566	.550
22	5	.549	.563	.650	.656	.607
23	6	.461	.335	.588	.604	.504
24	7	.437	.381	.429	.373	.405